**metin, yazı tipi, grafik, grafik tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**BYM308 YAPAY ZEKAYA GİRİİŞ**

**DÖNEM İÇİ PROJE RAPORU**

|  |  |
| --- | --- |
| **GRUP NO** | **9** |
| **GRUP ÜYELERİ** | **1) LALEGÜL YILDIRIM 210609016**  **2) FURKAN KARAN 210609010**  **3) ILGIN ÇITOK 210609011** |

İÇİNDEKİLER

1. TAKIM ŞEMASI………………………………………………………………………3
2. [2. PROBLEME EN YAKIN ÇÖZÜM SUNAN ULUSLARARASI MAKALELERİN ÖZETİ 3](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867064)-4
3. [3. KULLANILMASI PLANLANAN ALGORİTMALAR 4-5](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867065)
4. [4. KULLANILMASI PLANLANAN DONANIMLAR VE ÖZELLİKLERİ 5-6](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867066)
5. [5. ÖZGÜNLÜK 6](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867067)
6. [6. PROJE PLANI VE TAKVİMİ 6-7](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867068)
7. [7. REFERANSLAR 7](file:///C:\Users\Furkan%20Karan\Downloads\2024_SYZ-Kategori-1_Proje_Sunuş_Rapor_SablonuTR_V1jvQ%20(1).docx#_Toc155867069)

# TAKIM ŞEMASI

Bu projede üç takım üyesi olarak görev aldık ve her aşamada iş birliği içerisinde çalıştık. Tüm görevleri ortaklaşa gerçekleştirdik ve her birimizin projeye olan katkıları eşit derecedeydi.

**Görev Dağılımı ve Projeye Katkılar**

Veri Toplama ve Hazırlık: Tüm takım üyeleri olarak veri setini topladık ve yükledik. Verinin temizlenmesi, eksik ve aykırı değerlerin analiz edilmesi ve düzeltilmesi süreçlerini birlikte yürüttük. Bu aşamada, veri setinin anlamlı hale getirilmesi ve analize hazır hale getirilmesi için ortak çaba gösterdik.

Model Geliştirme ve Eğitim: Makine öğrenmesi modelinin oluşturulması, eğitilmesi ve optimize edilmesi görevlerini ortak bir şekilde gerçekleştirdik. Modelin doğruluğunu ve güvenilirliğini artırmak için birlikte çalıştık.

Sonuçların Görselleştirilmesi ve Raporlama: Projenin raporunu hazırlarken, sonuçların görselleştirilmesi ve sunulması konularında beraber çalıştık. Projenin dokümantasyonu, sonuçların anlaşılır hale getirilmesi ve nihai sunumun hazırlanması aşamalarında ekip olarak katkıda bulunduk. Her birimiz, raporun ve sunumun kalitesini artırmak için eşit derecede sorumluluk aldık.

# PROBLEME EN YAKIN ÇÖZÜM SUNAN ULUSLARARASI MAKALELERİN ÖZETİ

Bu bölümde, proje konumuz olan kalp hastalığı tahmini ile ilgili en yeni ve en ilgili üç uluslararası makaleyi özetledik. Her bir makale için problemi, çözümü, kullanılan veri setini ve elde edilen sonuçları belirttik.

**Makale 1: "Heart Disease Prediction Using Machine Learning Algorithms"**

Problem: Bu makale, kalp hastalığı riskinin erken teşhisi için makine öğrenmesi algoritmalarının kullanımını araştırmaktadır. Kalp hastalığı, dünya genelinde ölüm oranlarının yüksek olmasına sebep olan bir rahatsızlıktır ve erken teşhis hayati önem taşımaktadır.

Çözüm: Makale, çeşitli makine öğrenmesi algoritmalarını (lojistik regresyon, karar ağaçları, k-en yakın komşu, destek vektör makineleri) kullanarak kalp hastalığı tahmini yapmaktadır. En uygun modelin belirlenmesi için modellerin performansları karşılaştırılmıştır.

Veri Seti: Çalışmada Cleveland Heart Disease veri seti kullanılmıştır. Bu veri seti, yaş, cinsiyet, kan basıncı, kolesterol gibi çeşitli tıbbi özellikleri içermektedir.

Sonuçlar: En iyi sonuçları veren modelin lojistik regresyon olduğu bulunmuştur. Bu model, %85 doğruluk oranıyla kalp hastalığı tahmininde en etkili algoritma olarak belirlenmiştir.

**Makale 2: "Deep Learning Techniques for Heart Disease Diagnosis"**

Problem: Bu makale, kalp hastalığı teşhisi için derin öğrenme tekniklerinin uygulanabilirliğini araştırmaktadır. Derin öğrenme, büyük veri setleri üzerinde yüksek doğruluk oranları sağlayabilir, bu da tıbbi teşhislerde kritik bir avantajdır.

Çözüm: Makale, derin öğrenme algoritmaları (özellikle yapay sinir ağları ve derin sinir ağları) kullanarak kalp hastalığı teşhisi gerçekleştirmiştir. Modelin performansını artırmak için katman sayısı ve nöron sayısı gibi hiperparametreler optimize edilmiştir.

Veri Seti: Çalışmada Kaggle'dan alınan bir kalp hastalığı veri seti kullanılmıştır. Bu veri seti, çeşitli tıbbi özellikler ve teşhis bilgilerini içermektedir.

Sonuçlar: Derin öğrenme modelleri, %88 doğruluk oranı ile kalp hastalığı teşhisinde yüksek performans göstermiştir. Derin sinir ağları, diğer makine öğrenmesi algoritmalarına göre daha yüksek doğruluk sağlamıştır.

**Makale 3: "Hybrid Model for Heart Disease Prediction Using Ensemble Techniques"**

Problem: Bu makale, kalp hastalığı tahmini için topluluk (ensemble) yöntemlerinin etkinliğini incelemektedir. Farklı algoritmaların birleşimi, tek bir modelin performansını artırabilir.

Çözüm: Makale, birden fazla makine öğrenmesi algoritmasını birleştirerek hibrit bir model oluşturmuştur. Bu hibrit modelde, rastgele ormanlar (random forests), gradyan artırma (gradient boosting) ve AdaBoost gibi yöntemler kullanılmıştır.

Veri Seti: Çalışmada UCI Machine Learning Repository'den alınan bir kalp hastalığı veri seti kullanılmıştır. Veri seti, hastaların tıbbi geçmişi ve test sonuçları gibi bilgileri içermektedir.

Sonuçlar: Hibrit model, %90 doğruluk oranı ile en yüksek performansı göstermiştir. Bu sonuç, topluluk yöntemlerinin kalp hastalığı tahmininde etkili bir çözüm sunduğunu göstermektedir.

# KULLANILMASI PLANLANAN ALGORİTMALAR

Kullanılması Planlanan Algoritmalar:

Yapay Sinir Ağları (Artificial Neural Networks - ANN): Bu projede yapay sinir ağları kullanılması planlanmaktadır.

Neden Tercih Edildiği:

Projede yapay sinir ağlarının tercih edilmesinin nedenleri şunlardır:

Veri setinin karmaşıklığı: Projede kullanılan veri seti, kalp hastalığı riskini tahmin etmek için çeşitli özellikler içermektedir. Yapay sinir ağları, bu karmaşık özellikleri modelleme ve öğrenme yetenekleriyle öne çıkar.

Esnek mimari: Yapay sinir ağları, çok çeşitli mimari yapılarla uygulanabilir. Bu, modelin veri setine uygun bir şekilde özelleştirilebilmesini sağlar ve performansını artırabilir.

Derin öğrenme yetenekleri: Yapay sinir ağları, derin öğrenme prensiplerine dayanır. Bu da, veriler arasındaki karmaşık ilişkileri belirleme ve daha yüksek düzeyde özelliklerin otomatik olarak öğrenilmesini sağlama yeteneği ile projede etkili olmalarını sağlar.

Literatürdeki başarı: Yapay sinir ağları, kalp hastalığı riski tahmininde kullanılan modeller arasında literatürde yaygın olarak yer almaktadır. Önceki araştırmalar ve çalışmalar, yapay sinir ağlarının bu tür problemlerde iyi performans gösterdiğini göstermektedir.

Bu nedenlerle, projede yapay sinir ağlarının kullanılması planlanmıştır. Bu algoritmanın, projenin amacına uygunluğu ve geçmiş başarıları göz önüne alındığında, kalp hastalığı riskinin tahmininde etkili bir çözüm sunması beklenmektedir.

# KULLANILMASI PLANLANAN DONANIMLAR VE ÖZELLİKLERİ

Kullanılacak Araçlar ve Kütüphaneler

NumPy: NumPy, büyük çok boyutlu diziler ve matrislerle çalışmayı kolaylaştıran güçlü bir bilimsel hesaplama kütüphanesidir.

Neden Tercih Edildi: Veri manipülasyonu ve matematiksel işlemler için temel araç olarak kullanılır, hızlı ve verimli hesaplamalar yapar.

Pandas: Pandas, veri analizi ve manipülasyonu için güçlü araçlar sunar. Veri çerçeveleri ile çalışmak için idealdir.

Neden Tercih Edildi: Verileri temizleme, dönüştürme ve analiz etme süreçlerinde kullanılır, veriler üzerinde hızlı ve kolay işlem yapmayı sağlar.

TensorFlow: TensorFlow, Google tarafından geliştirilen açık kaynaklı bir derin öğrenme kütüphanesidir. Sinir ağları oluşturmak ve eğitmek için kapsamlı araçlar ve API'ler sunar.

Neden Tercih Edildi: TensorFlow, büyük ölçekli makine öğrenimi ve derin öğrenme uygulamaları geliştirmek için endüstri standardıdır. Kullanım kolaylığı, geniş topluluk desteği ve güçlü performansı ile tercih edilmektedir.

Matplotlib: Matplotlib, veri görselleştirme için kullanılan bir Python kütüphanesidir. Grafikler, çizelgeler ve diğer görsel öğeler oluşturmak için geniş bir araç seti sunar.

Neden Tercih Edildi: Model performansını ve verileri görselleştirmek için kullanılır. Eğitim sürecindeki metriklerin ve sonuçların görsel olarak analiz edilmesini sağlar.

Seaborn: Seaborn, Matplotlib üzerine inşa edilmiş bir veri görselleştirme kütüphanesidir ve daha estetik ve bilgilendirici grafikler oluşturmayı kolaylaştırır.

Neden Tercih Edildi: Veri görselleştirmelerinde daha gelişmiş ve estetik grafikler oluşturmak için kullanılır. Veri setinin keşfi ve sonuçların sunumu için idealdir.

Kullanılacak Veri Setleri, Kaggle platformundan temin edildi.

# ÖZGÜNLÜK

Projemizde kalp hastalığı riskini tahmin etmek için yapay sinir ağları kullanarak özgün bir yaklaşım benimsedik. Bu yaklaşımın özgünlüğü şu noktalarda yatıyor:

Özelleştirilmiş Mimari: Yapay sinir ağı mimarisi, karmaşık veri setimizin ihtiyaçlarına göre özelleştirilmiştir.

Derin Öğrenme Teknikleri: Derin öğrenme yöntemlerini kullanarak, diğer çözümlerden farklı bir yaklaşım sergiledik.

Yenilikçi Veri Ön İşleme: Veri ön işleme aşamasında geleneksel yöntemlerden farklı yaklaşımlar benimsedik.

Bu nedenlerle, projemiz mevcut çözümlerden farklı bir yaklaşım sunarak özgünlüğe katkı sağlıyor.

# PROJE PLANI VE TAKVİMİ

1. Proje hedeflerinin belirlenmesi ve gereksinimlerin toplanması (1 hafta)

2. - Veri setlerinin araştırılması ve seçimi (1 hafta)

- Veri temizleme ve ön işleme (1 hafta)

3. - Model eğitim verilerinin hazırlanması (1 hafta)

- İlk model prototipinin geliştirilmesi (1 hafta)

4. - Modelin eğitimi (2 hafta)

- Modelin doğrulama ve test aşaması (1 hafta)

- Model optimizasyonu (1 hafta)

5. - Uygulamanın test edilmesi (1 hafta)

6. - Sonuçların analiz edilmesi ve raporlanması (1 hafta)

- Nihai raporun yazılması ve sunumu (1 hafta)

Detaylı Görev Tanımları

- Proje Tanımlama: Projenin ana hedeflerinin ve başarı kriterlerinin belirlenmesi. Gereksinim analizinin yapılması.

- Veri Toplama ve İşleme: Kalp hastalıklarıyla ilgili mevcut veri setlerinin araştırılması, seçilmesi ve gerekli ön işlemlerin yapılması.

- Model Geliştirme: TensorFlow kullanılarak model mimarisinin tasarlanması, eğitim verilerinin hazırlanması ve model prototipinin geliştirilmesi.

- Model Eğitimi ve Değerlendirme: Modelin eğitim süreçlerinin gerçekleştirilmesi, doğrulama ve test aşamalarının yapılması, optimizasyon süreçlerinin tamamlanması.

- Test ve Validasyon: Uygulamanın kapsamlı testlerinin yapılması.

- Sonuçların Raporlanması: Sonuçların analiz edilmesi, raporlanması ve nihai raporun hazırlanarak sunulması.

# REFERANSLAR

1.) https://www.sciencedirect.com/science/article/pii/S1877050920310195

2.) https://www.mdpi.com/2075-4418/12/12/3215

3.) https://curiousily.com/posts/heart-disease-prediction-in-tensorflow-2/

4.) https://jeas.springeropen.com/articles/10.1186/s44147-023-00280-y

5.) https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03626-y

6.) https://dergipark.org.tr/tr/pub/bbd/issue/49546/508902

7.) https://dergipark.org.tr/tr/pub/akufemubid/issue/71191/1114346